

# Cloning and Characterization of 3.1kb Promoter Region of the Oct4 Gene from the Fischer 344 Rat

Hong He, Mark McHaney, James Hong and Mark L. Weiss\*

Department of Anatomy and Physiology and the Midwest Institute for Comparative Stem Cell Biology, Kansas State University, Manhattan, KS 66506, USA

**Abstract:** Here, the role of methylation in regulation of rat Oct4 gene was evaluated during embryonic development, in adult tissues and in embryo-derived cells. First, the region 3.1 kb upstream to the rat Oct4 ATG site was cloned and sequenced. The rat Oct4 upstream sequence was similar to that in bovine, mouse and human with two upstream elements: proximal (PE) and distal enhancers (DE) and four homology conserved regions (CR1-4). The conserved regions in the rat have 69% - 96% homology with bovine, human, mouse sequences. Next, the methylation pattern in the promoter was determined during embryonic development, in adult tissues, in rat embryonic stem cell (ESC)-like cells and umbilical cord-derived cells (the feeder for ESC-like cells) using the bisulfite method and DNA sequencing. The promoter was methylated in adult and fetal tissues, and in days post coitus (DPC) 10.5 and 12.5 embryos and hypomethylated in DPC4.5 embryos and in rat ESC-like cells. The expression of Oct4 was evaluated by qRT-PCR. DPC4.5 embryos and rat ESC-like cells had higher expression of the Oct4 gene compared to DPC10.5 and 12.5 embryos, adult tissues and embryoid bodies derived from rat ESC-like cells. Thus, the methylation status correlated with the qRT-PCR results. These results indicate that the rat Oct4 3.1kb promoter region is organized and contains transcription binding and regulatory sites similar to those described for bovine, mouse and human. The rat Oct4 promoter is methylated during embryonic development after 4.5 DPC and during differentiation of rat ESC-like cells to embryoid bodies.

## INTRODUCTION

By the cloning of various mammalian somatic cell nuclei and by the fusion of somatic cells with embryonic stem cells (ESCs), it has been demonstrated that the nuclei of differentiated somatic cells are genetically equivalent to zygotic nuclei and that somatic cells can be reprogrammed into pluripotent cells [1-3]. The recent studies from Tanaka's, Jaenisch's and Thomson's labs show that the adult somatic cells can be reprogrammed back into embryonic-like pluripotent stem cells *in vitro* and indicate that expression of a few transcription factors are enough to induce pluripotency in adult somatic cells, e.g., reprogram the cells to embryonic-like pluripotent stem cells (e.g., [1-7]). The *Pou* family member, Oct4 is one of the few transcription factors which can reprogram adult somatic cells into pluripotent stem cells.

The Oct4 (*Pou5f1*) gene encodes a transcription factor and is specifically expressed in the inner cell mass (ICM), ESCs and germ cells at high levels [8]. Most cells lose Oct4 expression as they differentiate into somatic tissues and Oct4-deficient embryos and ESCs spontaneously differentiate into trophectoderm cells [9, 10]. The expression of Oct4 is essential for ESCs to establish and maintain pluripotency; and understanding the functional regulation of Oct4 gene expression is central to understanding pluripotency and the downstream developmental program.

One major modification of genes and the genome is DNA methylation which suppresses expression of methylated genes [11-13]. In mouse, the promoter of Oct4 gene is distinctly methylated in somatic cells and is not expressed. In contrast, the promoter of Oct4 is unmethylated in undifferentiated ESCs and embryonal carcinoma (EC) cells which express the Oct4 gene [14-16]. Therefore, methylation of the Oct4 promoter is one important mechanism of Oct4 gene expression regulation. Here, we cloned the Fischer 344 (F344) rat Oct4 3.1 kb upstream to the ATG and characterized this region during embryonic development by analyzing the methylation pattern of the Oct4 promoter at -217, -207, -165, -113, -43, -24, and -19 bp from the translation start site (designated as +1) in rat embryos, and contrasted that pattern with that seen in adult cells from liver and testes, rat umbilical cord mesenchymal stromal cells, rat ESC-like cells (e.g. cells derived from the blastocyst inner cell mass and maintain *in vitro* for 3-9 passages). Finally, quantitative RT-PCR was used to evaluate Oct4 expression in these tissues and in embryoid bodies derived from rat ESC-like cells. The results indicate that the regulatory elements in the 3.1 kb upstream from the ATG include a proximal promoter region and a proximal and distal enhancer region. Within these elements are four conserved regions (CR). CR1, 2 and 4 of the rat contains putative transcription factor and hormone binding sites that are likely to play a key role in regulation of Oct4 transcription. Next, the methylation pattern of the F344 rat Oct4 promoter was evaluated during development days post coitus (DPC) 4.5-12.5 and in adult tissues, in umbilical cord mesenchymal stromal cells, and rat embryonic-like stem cells (ESCs). The promoter was methylated in DPC 10.5, DPC

\*Address correspondence to this author at the Department of Anatomy and Physiology, Kansas State University, Manhattan, KS 66506, USA; Tel: 785-532-4520; Fax: 785-532-4557; E-mail: Weiss@Vet.Ksu.Edu

12.5 and adult tissues and in umbilical cord derived tissues, but not in rat ESCs or in DPC 4.5 embryos. The methylation of the promoter paralleled inversely with the Oct4 gene expression determined by qRT-PCR. Specifically, higher levels of gene expression was observed in DPC 4.5 and ESC cells at passage 3 and 9 than in adult tissues, umbilical cord tissues, DPC 10.5 and 12.5 and in rat ESC-like cells differentiated to embryoid bodies.

## MATERIALS AND METHODS

### Cloning DNA of Oct4 Upstream to the ATG

The rat mRNA sequence of rat Oct4 was obtained from Genbank (accession no. NM\_001009178) and blasted to the rat genome. Next, 3.1kb of rat genomic DNA upstream to the Oct4 ATG was obtained from Genbank and from this we generated primers (Forward: 5'-GGG GAA GGT GGG CTC CCC GA-3'; Reverse: 5'-CCT GGG GAC TCT GAT GGT TA-3'). Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega) from five Fischer 344 rats (IACUC approval was obtained for all animal work). To amplify genomic DNA, PCR was performed under the following condition: the initial denaturation of 5 min at 95 °C; 35 cycles of 30 sec at 94°C, 30 sec at 54 °C, 4 min at 68 °C and the final extension of 30 min at 68 °C. High fidelity Taq was used for this PCR (Invitrogen). PCR amplified a band of 3.1 kb. This DNA was isolated from the agarose gel, purified and cloned using a TOPO TA cloning kit (Invitrogen). Nine positive clones were sent for sequencing. After sequencing, the sequences of the nine clones were 100% identity.

### DNA Methylation

#### Collection of Embryos and Culture Cells

Time-pregnant Fischer 344 (F344) rats were used to obtain embryos (coitus was determined by sperm in vaginal lavage fluid following overnight housing with proven studs; conception was assumed to take place at midnight, DPC 0). Embryos were collected on day 4 after conception (D4.5), D8.5, D10.5, D12.5. Embryo-derived cells were derived from the inner cell mass from D4.5 embryos using standard methods with modification (details available from the authors). Adult liver, brain and testes tissues were collected from adult F344 following humane sacrifice. Rat umbilical cord cells were collected following human sacrifice of time pregnant dams at D19.

#### Bisulfite DNA Methylation Analysis:

The extraction of genomic DNA and bisulfite mutagenesis sequencing analysis were conducted by Wizard Genomic DNA Purification Kit (Promega, Madison, WI) and EZ DNA Methylation Kit (Zymo Research, Orange, CA) respectively [17]. The mutagenized DNAs were amplified by PCR as follows: the initial denaturation at 95°C for 5min, 35 cycles of [94°C for 1min, 53°C for 30sec, and 72°C for 30sec, and the final extension at 72°C for 3min]. The PCR products were run on an agarose gel and the amplified DNA bands were purified by Gel Extraction Kit (QIAGEN). The purified PCR products were ligated into the TOPO TA plasmid (Invitrogen) for sequencing. The PCR primers used were for methylation analysis were: Forward 5'-GTGATGGGGATTT AAGTAATTGGTTTT-3'; Reverse 5'-TCTAAAACCAAT

ATCCAACCATAA-3'. This primer produced a 375 bp fragment referenced to Genbank number: EU419996 as 2710-3085.

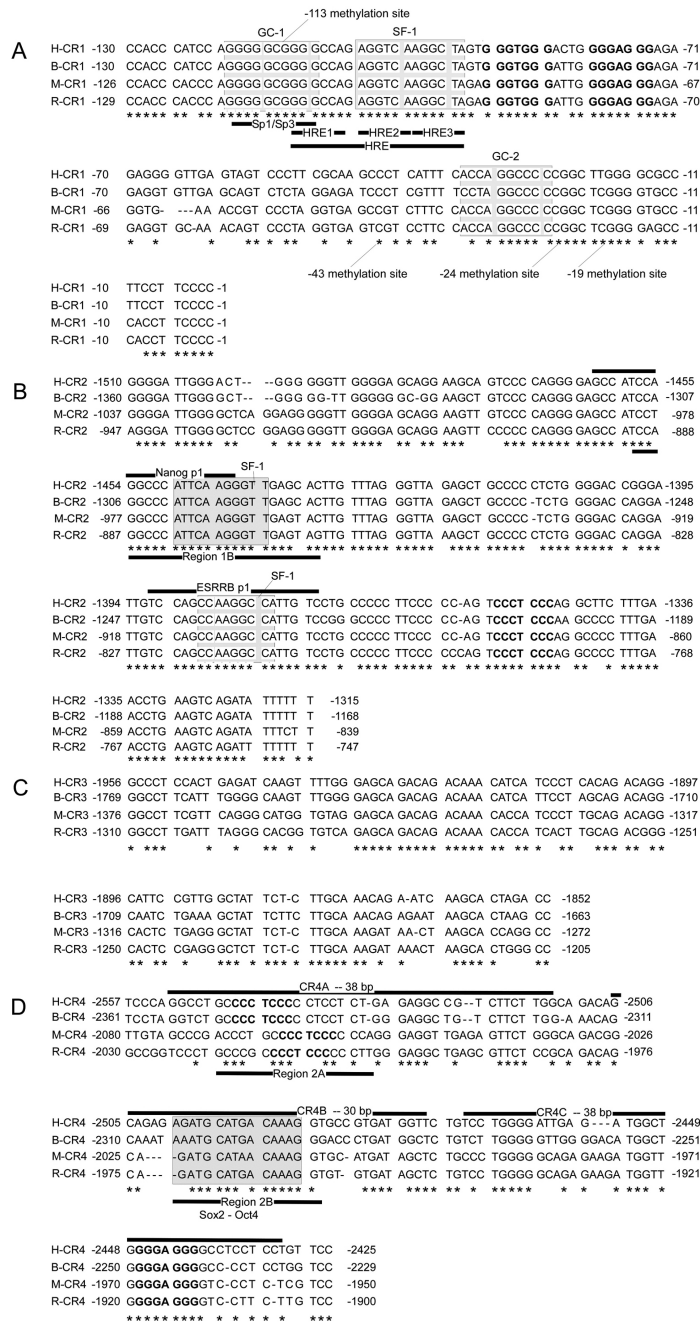
### Quantitative Reverse Transcription- Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated by RNeasy micro kit using manufacture's protocol (QIAGEN, cat # 74004) with on-column DNase I digestion from adult rat liver, brain, testes, rat embryos at D4.5, D8.5, D10.5, D12.5 of development, rat embryonic stem cell-like cells (ESC-like, passage 3, 6, 12) and rat umbilical cord mesenchymal stromal cells (passages 1-7, the mitotically-inactivated feeder layer (data from a few passages is shown), and from rat ESC-like cells that were plated on low-adherence dishes in feeder-free conditions without cytokines for 1 and 2 weeks and allowed to form embryoid bodies (EB). Total RNA was reverse transcribed using Superscript III first strand synthesis system for RT-PCR (Invitrogen, cat # 18080-051). The concentration and quality of the cDNA (and the genomic DNA contamination) was evaluated by Nanodrop and by PCR for porphobilinogen deaminase (PBGD, also called hydroxymethylbilane synthase), a low-abundance housekeeping gene, using a primer pair that spans an intron. The qRT-PCR was done in duplicate and at two concentrations of starting total RNA using Sybr/ Fluorescein Green PCR Master Mix (SABiosciences, cat # PA-011) on a Bio-Rad iCycler (program: 95°C for 15 minutes, followed by 45 cycles of 95°C for 15 sec, 56°C for 15 sec, and 60°C for 1 min). Following the qRT-PCR program, the products were subjected to melting point analysis to confirm amplification of a single product. The real time primers for rat PBGD (163 bp) Forward- 5'- GCA CGG CAG CTT AAT GAT GT-3'; reverse- 5'-CAA GGC CGA AGT CTC AAC AC-3'; rat Oct4 (132 bp) forward- 5'-AGA ACC GTG TGA GGT GGA AC-3'; reverse- 5'-GCC GGT TAC AGA ACC ACA CT-3'.

## RESULTS

### Comparison of the Sequences

The 3.1kb DNA sequence upstream to the ATG was sent to Genbank (Genbank accession no. EU419996). The rat Oct4 gene upstream promoter sequence, like mouse, human and bovine, has four conserved regions (CR1-4). Fig. (1) shows the alignment of the four conserved regions (CR1, CR2, CR3 and CR4) in the human, bovine, mouse and rat Oct4 gene orthologs. CRs have high GC content and high homology between the four species, from 70% to 96% (Table 1). As shown in Fig. (1), CR1 corresponds to nucleotides -129/-1 (F344 rat); CR2 to -947/-747; CR3 to -1310/-1205 and CR4 to -2030/-1900 respectively. The putative Sp1/Sp3 binding site (5'-GGGGCGGGG-3') and HRE regions 1-3 show 100% homology in all four species, as does the putative SF-1 site. In CR2, Region 1B of the proximal enhancer was identified [18], as was a putative Nanog, E-box/Mash-2 and ESSRB binding site [19]. Two putative SF-1 binding sites were identified [20]. CR3 had a central region that was conserved -1285/-1272). In CR4, Region 2A and 2B of the distal enhancer [21]. A schematic map of the four conserved regions (CR) of the Oct4 5' upstream region from the human, bovine, mouse and rat orthologs is depicted in Fig. (2A) (Figure modified from [22]).



**Fig. (1).** Alignment of the rat (R) Oct-4 proximal promoter 5' upstream sequences with its human (H), bovine (B) and mouse (M) orthologs. (A) Alignment of the conserved region 1 (CR1), or proximal Oct-4 promoter. The putative Sp1/Sp3 binding site (5'-GGGGGCGGGG-3') [41, 59] (rat -118/-109) has been bolded and gray, and the overlapping HRE (5'-GGGCCAGAGGTC AAGGCTA-3'; [41]; rat -111/-93) has been underlined. Both the Sp1/Sp3 site and the HRE are conserved in the human, bovine, mouse and rat Oct-4 5' upstream regions. The R1, R2 and R3 regions of the HRE are located at rat sequences -111/-106, -104/-99 and -98/-93, respectively are indicated (HRE1-HRE3 as noted). DR1 is located between the R1 and R2 (rat -105). Two GC rich regions have been boxed in gray boxes. The putative methylation site at -113 within Sp1/Sp3 and GC-1 has been indicated. Two GGG(T/A)GGG domains 3' to HRE are indicated by bolded text. The SF-1 binding site within HRE is indicated by a gray box. (B) Alignment of the conserved region 2 (CR2). Region 1B (5'-TCCAGGCCCATTC AAGGG TTGAGTAG-3'; [18], rat -892/-866) of the proximal enhancer (PE) is indicated by the line underneath. Within region 1B, a SF-1 binding site is indicated by a gray box. The putative Nanog binding site is indicated overlapping region 1B by the line on top. Overlapping the 3' border of Region 1b is a putative E-box/Mash2 binding site (indicated by gray box). The ESRRB putative binding site is indicated overlapping with a second SF-1 site. A CCTCCC motif is indicated by bolded text. (C) Alignment of the conserved region 3 (CR3), and (D) the conserved region 4 (CR4). Region 2A (5'-GCCCCGCCCTCCCCCTT-3') [41] (rat -2020/-2003) of the distal enhancer (DE) has been underlined in CR4A. The putative Oct4 Sox2 binding domain is indicated with Region 2b by the gray box. Regions CR4a-c are indicated by the line on top. A conserved GGGAGGG motif is indicated by bold text. The bovine sequences are from [60] (Genbank accession no. AF022986). The human and mouse promoter sequences from [61] (Genbank accession no. Z11900) and [62] (Genbank accession no. AJ297528). The rat sequences are from Genbank accession no. EU419996. (Figure modified from reference [22]). Nucleotides have been numbered relative to the translational start site and are indicated at the beginning of the alignment, and at the right of each sequence. Asterisk denotes nucleotides that are identical between the four species. Dashes between nucleotides serve to align the sequences.

**Table 1. Homology between Rat, Human, Mouse, and Bovine Conserved Regions**

Conserved Region (CR)	r-h (%) Equal bp/Total bp	r-m (%) Equal bp/Total bp	r-b (%) Equal bp/Total bp	h-m (%) Equal bp/Total bp	h-b (%) Equal bp/Total bp	m-b (%) Equal bp/Total bp
CR1	84%	96%	86%	84%	88%	86%
CR2	92%	95%	93%	94%	95%	95%
CR3	75%	88%	77%	77%	80%	81%
CR4	77%	80%	70%	71%	86%	70%

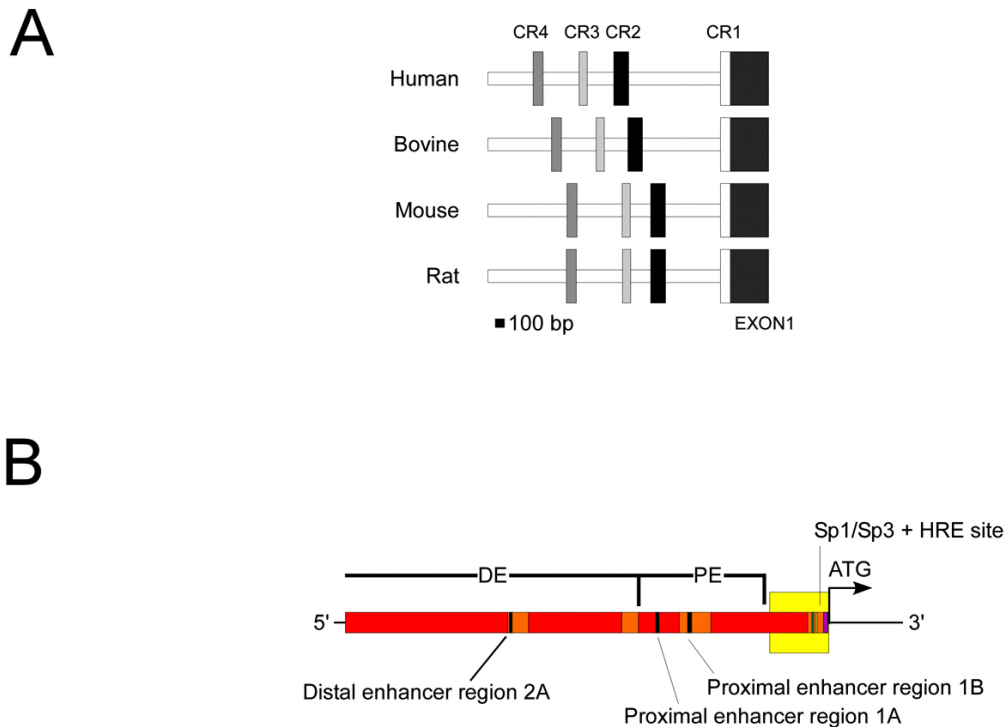
r refers to rat sequences, h refers to human sequences, m refers to mouse sequences, b refers to bovine sequences.

### Methylation Sites of the Oct4 Gene

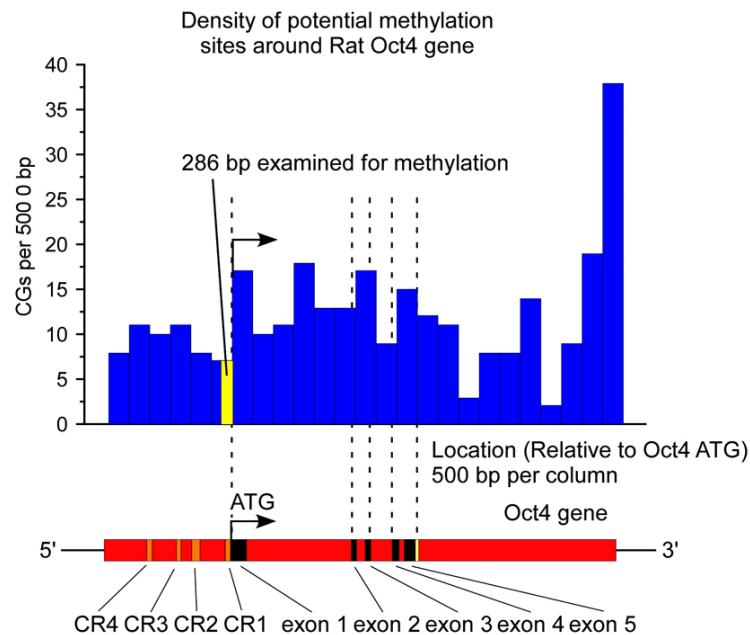
As schematized in Fig. (2B), the Oct4 gene has two distinct enhancers, a distal enhancer (DE) and a proximal enhancer (PE), which drive the cell type-specific expression of the Oct4 gene [23, 24], upstream from the proximal (minimal) promoter. An analysis of the CpG sites from 3.1 kb 5' to the ATG through the exons and introns of the Oct4 gene reveals that more potential methylation sites fall within exons 1-5 and intron 1 than in the proximal promoter (see Fig. 3). In other species, there are more CpG islands in the PE which means that it is important for regulation of Oct4 gene expression [15, 17].

### Methylation of the Minimal Promoter of Oct4 during Development, in Adult Tissues and in Rat ESC-Like Cells

Here, we investigated the methylation status of CpG of the region 5' to the Oct4 gene ATG containing the proximal promoter region (-217, -207, -165, -113, -43, -24, -19 bp from the translation start site (designated as +1) by sodium bisulfate genomic sequencing in rat embryo-derived cells (called ESC-like cells here), D4.5, D10.5, D12.5 embryos (Fig. 4A) and F344 rat adult liver (Fig. 4B). In two independent trials, ten clones of each sample were sequenced. The 7 CpGs in this region comprise all the CpGs in the F344



**Fig. (2).** (A) Schematic representation of the *Oct-4* 5' proximal promoter upstream region in the rat, human, bovine, and mouse orthologs, denoting the relative size and location of each of the conserved regions (CR). CR 1 is represented by a white square, CR 2 by a black square, CR 3 by a light gray square, and CR 4 by a dark gray square. CR 1 is located at -129/-1 in the rat sequence and is located at -126/-1 in the mouse sequence and at -130/-1 in the human and bovine sequences. CR 2 is located at -947/-747, -1037/-839, -1360/-1168 and -1510/-1315; CR 3 is located at -1310/-1205, -1376/-1272, -1769/-1663 and -1956/-1852 1956/-1852, and CR 4 is located at -2030/-1900, -2080/-1950, -2361/-2229, and -2557/-2425 in the rat, mouse, bovine, and human sequences, respectively. An arrow denotes the direction of *Oct-4* gene transcription. (B) Schematic representation of the rat *Oct-4* 5' proximal promoter upstream region denoting the location of specific distal enhancer (DE) and proximal enhancer (PE) sites with respect to the CRs. CR 1, CR 2, CR 3, and CR 4 are located at sequences -129/-1, -947/-747, -1310/-1205, and -2557/-1900, respectively. Site 2A [41] of the DE is located at sequence -2020/-2003 in CR 4; site 1A [18] of the PE is located at sequence -1094/-1073 between CR 3 and CR 2. An arrow denotes the direction of *Oct-4* gene transcription and the ATG location. The yellow box indicates the region analyzed for methylation. (Figure modified from [22]).



**Fig. (3).** Distribution of CG sites from -3000 to +9500 (ATG = +1) of the Oct4 gene. Top: The number of CG sites in each 500 bp is shown. Bottom: Drawn to scale, a schema of the important elements of the Oct4 gene. The dotted lines indicate the corresponding methylation in exon 1, and exons 2-3 and exons 4-5. Note the relatively higher numbers of CG sites found in exons 1, exons 2-3 and exons 4-5 and that intron 1 has a region of high number of CG sites. Note that the proximal promoter region (yellow bar 5' to the Oct4 ATG) does not have a large number of CG sites.

rat Oct4 minimal promoter, 4 of the CpGs are found in CR1 (indicated in Fig. 1A). In D4.5 embryos, very low levels of methylation were detected (0% in one trial and 1% in the other). In rat ESC-like cells, low levels of methylation were detected, too (1% in one trial and 9% in the second trial, see Fig. 4B). Thus, undifferentiated rat ESC-like cells are hypomethylated (comparing Fig. 4A and Fig. 4B). In D10.5, more methylation was found (26% of the CpG sites in one trial and 51% in the other). In D12.5, higher levels of methylation were detected (60% of the CpG sites in one trial and 64% in the other trial). The methylation in adult tissues was similar to that of D12.5 embryos: adult liver (67% of the CpG sites methylated in one trial and 59% in the second, Fig. 4B and 4D), adult testis (54% methylated, Fig. 4D). The methylation status of F344 rat umbilical cord mesenchymal stromal cells (feeder layer for the rat ESC-like cells) was evaluated at P1 through P7, Fig. 4C). The umbilical cord tissue was methylated like D12.5 embryos and adult tissues (63-84% methylated in the different trials). The Oct4 promoter region is methylated during development, mostly at sites -40, -113, -165, -207 and -217 (5' to the start site). Note that methylation at site -113 lies within the Sp1 binding site. In data not shown, the methylation status of Oct4 in bone marrow mesenchymal stromal cells was evaluated and was not different from other adult tissues. A box and whisker plot of the methylation data is shown in Fig. (5). As shown in Fig. (5), the promoter of Oct4 is hypomethylated in rat ESC-like cells and D4.5 embryos, relative to adult liver and D10.5 and D12.5 embryos and rat umbilical cord mesenchymal stromal cells (RUC).

### qRT-PCR Results

To independently evaluate the methylation status of the Oct4 proximal promoter, expression of the Oct4 gene was determined from total RNA was collected from D4.5, D10.5,

D12.5, adult liver and adult brain, blastocyst-derived cells at passage 3 and passage 9, and embryoid bodies after 7 days and 14 days of differentiation. These results are shown in Fig. (6). Oct4 and PBGD were run in duplicate at two different concentrations of total RNA (4 and 40 ng). The Ct values for the PBGD (a low-abundance housekeeping gene used to standardize the data) averaged  $26.6 \pm 0.2$  (average  $\pm$  SEM; range 24.35 --33.7) for 4 ng starting total RNA and  $22.9 \pm 0.6$  (range 20.5 – 30.5) for 40 ng. The D4.5 embryos expressed Oct4 at relatively high levels, indicated by the delta Ct value of -2.9 and -3.15 (for 4 and 40 ng starting total RNA). In contrast, Oct4 expression was lower in D10.5 and D12.5 embryos indicated by delta Ct values in the 6.2 -- 8.85 range (for 4 and 40 ng starting total RNA). The delta Ct values for Oct4 for adult brain ranged from 2.9 – 7.55 and for adult liver ranged from 5.2 – 11.65. These values suggest that Oct4 expression in adult liver and brain was similar to the expression level in D10.5 and D12.5 embryos. The rat blastocyst-derived cells were evaluated at passage 3 and 9. The delta Ct value of the embryonic stem cell-like cells was -2.15 – 2.8 and there was a trend for the earlier passage cells to have higher expression than the passage 9 cells. Thus, the Oct4 expression of the ESC-like cells was similar to D4.5 embryos. Following differentiation of ESC-like cells to embryoid bodies for 7 days or 14 days, the delta Ct value was 4.85 – 8.45. There was a trend for the embryoid bodies that had been differentiated for 7 days to have higher expression than the embryoid bodies differentiated for 14 days. The Oct4 delta Ct values for both 7 and 14 day differentiated embryoid bodies was similar to that observed for D10.5, D12.5 embryos and for adult liver and brain.

### DISCUSSION

Oct4 is a member of the POU homeodomain family of transcription factors and is expressed during preimplantation

embryonic development and in pluripotent stem cells, germ cells and embryo carcinoma cells [25]. Oct4 is very important for the maintenance of pluripotency [10] and overexpression of Oct4 and Klf4 in combination with small molecule compounds is sufficient to induce pluripotent stem cells to form from mouse embryonic fibroblasts [26] or Oct4 and Sox2 in the case of human fibroblasts [27]. Oct4 acts as a transcriptional regulator that activates genes that are critical to maintain the pluripotent state and represses genes that induce differentiation [28]. Overexpression of the Oct4 gene in mouse embryonic stem cells results in the primitive endoderm and downregulation of Oct4 results in ESC differentiation [10]. Differences are found between rat and mouse epiblasts [29], and it has been difficult to initiate rat embryonic stem cells and maintain them in the undifferentiated state [30-35]; problems with initiating and maintaining rat ESCs are no longer relevant [36-38]. For these reasons, we sought to evaluate whether differences would be found in promoter and enhancers upstream of the Oct4 gene in the rat, human and mouse. Next, since methylation of the Oct4 proximal promoter is a epigenetic regulator of the during development [15, 39] in human and mouse cells, we sought to evaluate the methylation status of the rat Oct4 proximal promoter region during embryonic development, in rat embryo-derived stem cells and in adult tissues. We showed a high degree of simi-

larity of the organization of the rat Oct4 upstream elements including proximal and distal enhancer elements and the minimal promoter, to mouse, bovine and human. In addition, we found a parallel between the methylation status of the minimal promoter and Oct4 gene expression determined by qRT-PCR. Finally, the downregulation of the rat Oct4 gene during embryonic development follows a similar time course to what has been described in mouse.

Oct4 gene expression during development is orchestrated by a complex arrangement of regulatory elements. The protein levels of the master regulators, Oct4, Nanog and Sox2, are regulated by both autoregulatory and feed-forward transcriptional mechanisms in a program that determines the timing of differentiation. Oct4, Nanog and Sox2 modulate the expression of Oct4 by binding to and activating the Oct4 promoter [40]. In addition, previous studies have found, in mouse ESCs and ECs, that Oct4 expression is controlled by two separate enhancer elements in the upstream to a TATA-less minimal promoter [23, 41]. The distal enhancer is active in the undifferentiated ESCs, the inner cell mass (ICM) and primordial germ cells [23], and is the site of the putative Oct4 Sox2 binding site [42]. The proximal enhancer is active in EC and primitive ectoderm and contains retinoic acid response elements that downregulate Oct4 expression [18]. The proximal enhancer is also the site of Nanog binding site

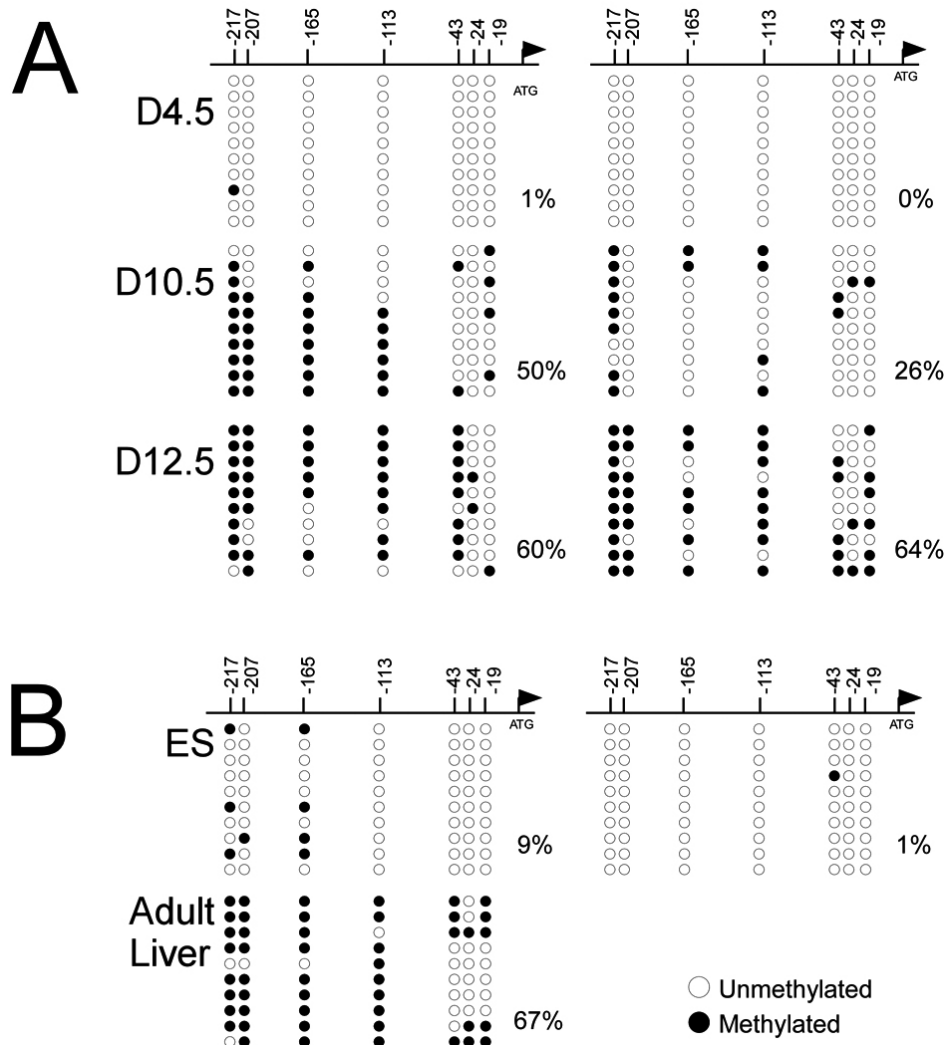
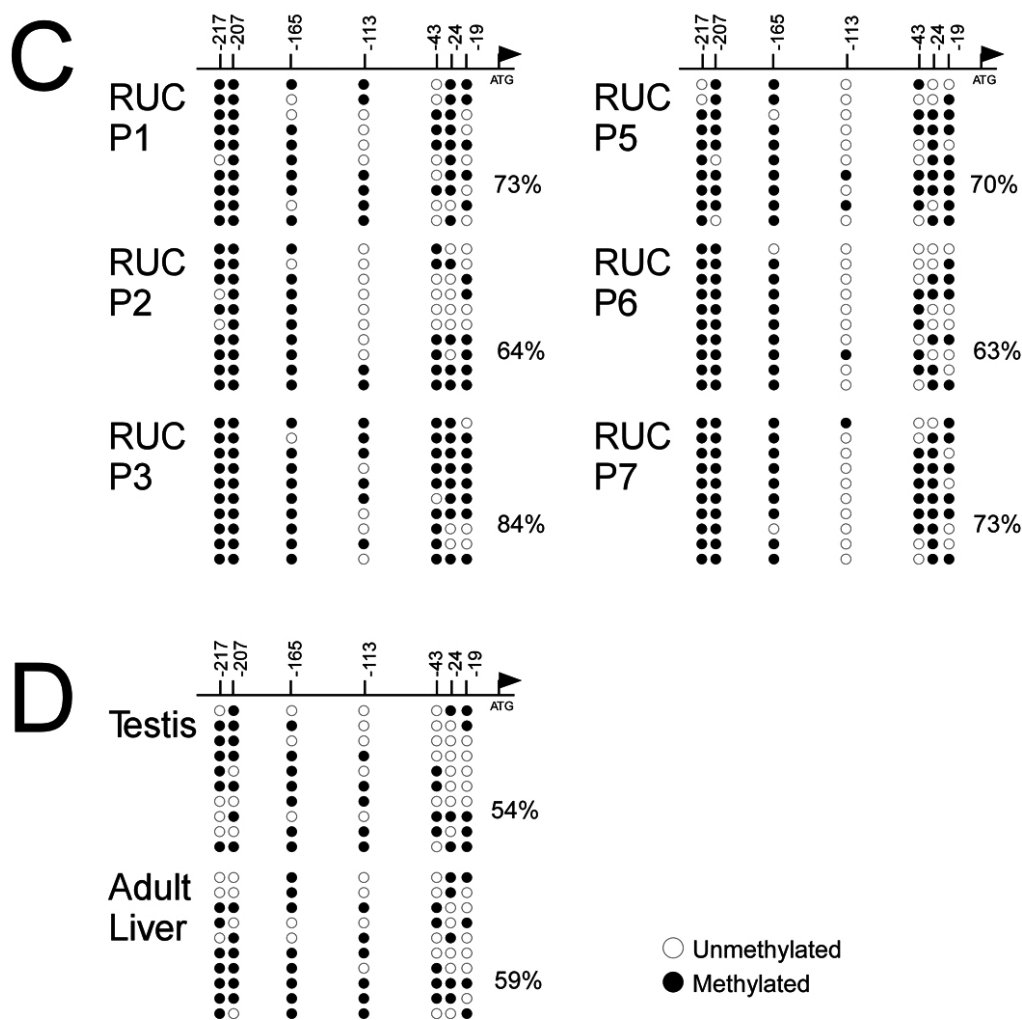


Fig. (4) contd....



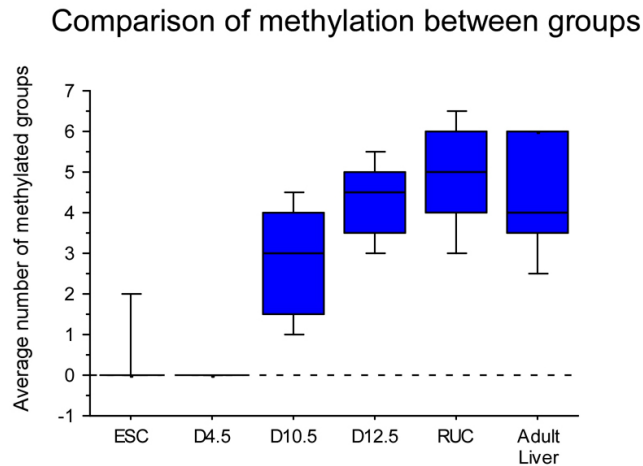
**Fig. (4).** DNA methylation status of the Fischer 344 (F344) rat Oct4 proximal promoter region. DNA methylation profile of the individual CpG sites in the Oct4 upstream region (-19 to -217) in (A) F344 rat embryos at D4.5, D10.5, D12.5, and (B) F344 rat embryonic stem cell-like cells and F344 adult rat liver and (C) in F344 rat umbilical cord mesenchymal stromal cells (RUC) at P1-3, left and P5-7 right, and (D) F344 rat adult testes and liver by sodium bisulfate sequencing. The map at the top indicates the CpG dinucleotide location within the Oct4 proximal promoter region relative to the start site. Each PCR product was subcloned and 10 clones sequenced. Except for the adult testes, the results from two independent trials are shown (independent genomic DNA samples used). The methylation status, either unmethylated (open circle) or methylated (closed circle), is indicated at each CpG site. Numbers to the right indicate the average percentage of methylated sites for each trial of 10 clones.

and estrogen response elements [19], steroid-thyroid response elements.

Within the 3.1kb upstream of the Oct4 ATG, four conserved regions (CR) have been identified [22]. Three of these regions, CR1, CR3 and CR4 are very important for the regulation of Oct4 gene expression. In CR1 (-129 -- -1 in the rat), four important elements are found: the Sp1/ Sp3 transcription factor binding domain, the hormone response element (HRE), and two GC rich domains (GC-1, -119 -- -113; GC-2, -34 -- -24). GC-1 and the Sp1/Sp3 and HRE are conserved between the four species. GC-2 is conserved between rat, mouse and human. The Sp1 transcription factor is thought to be important in the regulation of transcription in TATA-less minimal promoters, such as Oct4 [43]. While the Sp1 site is able to bind factors in both untreated and retinoic acid (RA)-treated EC, the HRE binds nuclear receptor superfamily members differentially in cells before and after RA treat-

ment. Thus, the HRE is thought to have both positive and negative roles in Oct4 expression [41, 44-46]. Again in CR1, proximal promoter activity is almost completely lost by deletion of GC-1 (which contains the Sp1 site) and virtually unaffected by GC-2 deletion [47]. As noted previously by Nordhoff, two GGG(T/A)GGG elements are located 3' to the HRE/SF-1 domain in CR1 (-90 -- -84, -79 -- -74), these domains are conserved and their function is unknown [22].

CR-2 contains region 1B of Okazawa [18]. Within and adjacent to Region 1B is a putative Nanog binding site (-895 -- -874) and a putative binding site for estrogen related receptor  $\beta$  (ESRRB p1, -824 -- -806) [19]. Also within CR2 are two putative SF-1 binding sites, one is located within region 1B (-882 -- -871) and the other overlaps with the ESRRB binding site (-818 -- -810) [20]. A CCCTCCC motif was conserved, as noted previously in mouse, human and bovine [22].

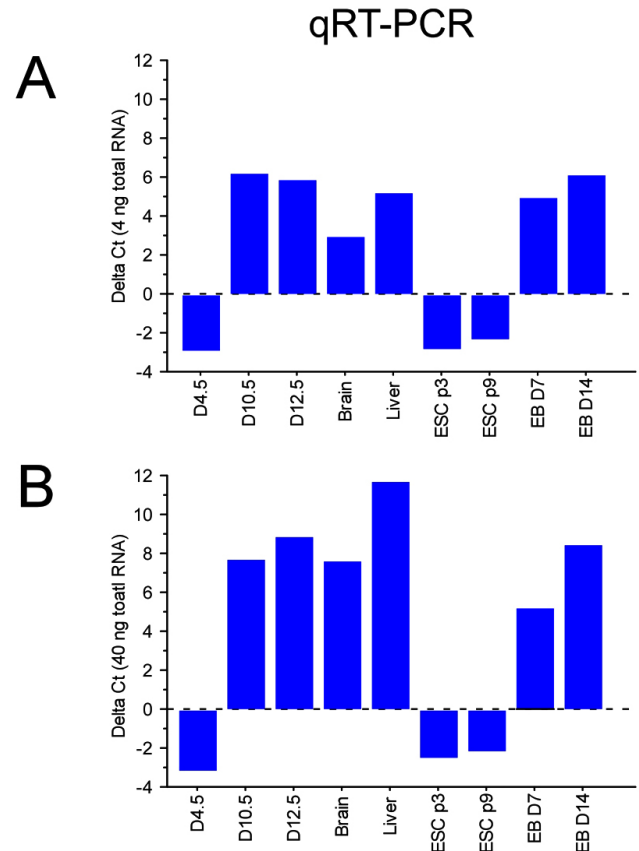


**Fig. (5).** Box and whisker plot showing distribution of methylation in the different experimental groups. Box and whisker plot indicates 90<sup>th</sup> percentile, 75<sup>th</sup> percentile, median, 25<sup>th</sup> percentile and 10<sup>th</sup> percentile for each distribution. Abbreviations: ESC, rat embryonic stem-like cells; D4.5, embryos collected at 4.5 days post coitus; D10.5, embryos collected at 10.5 days post coitus; D12.5, embryos collected 12.5 days post coitus; RUC, F344 rat umbilical cord mesenchymal stromal cells.

CR-3 lies within the distal enhancer. A conserved region was noted 5'-GAG CAGACAGACAAACACCAT-3' (-1285/-1277). CR-4 contains the CR4a-c regions. Within the CR4b is region 2B of Okazawa [18] and a putative Oct4 Sox2 binding region (5'-GATGCATGACAAAG-3') [42]. In addition, CR4 contains CCC(T/A)CCC motifs in Region 2a (note that the domain is conserved in human and bovine and in mouse and rat with a shift) and 3' to region 2b in CR4c, as previously noted in mouse, bovine and human [22].

Review of the location of potential methylation sites indicated that many CG sites were located with exons 1, 2-3 and 4-5 and within intron 1. Previous studies had shown that methylation of the proximal promoter region and proximal enhancer correlated well with differentiation and with decreased expression of Oct4 [15, 39]. Furthermore, during differentiation of human teratocarcinoma cells by retinoic acid, methylation occurred in parallel in the proximal promoter region and in exon 1. The methylation in proximal promoter and exon 1 paralleled the suppression of Oct4 gene expression [39]. Deb-Rinker *et al.* suggested that methylation is unlikely to play a direct role in downregulation of expression, but rather "locks in" the stable, silenced state [39]. Clearly, the lock can be "picked" by transient overexpression of certain transcription factors and induction of pluripotency. Similarly, recent work with rat embryonic stem cells has shown that the Oct4 proximal promoter and the Nanog proximal promoter are hypomethylated in undifferentiated cells [38]. This corresponds with what was reported here for the methylation status of the proximal promoter of Oct4 in the F344 rat ESC-like cells and F344 D4.5 embryos. In addition, the proximal promoter region of Oct4 gene is hypomethylated in rat D4.5 embryos and is methylated in D10.5 and D12.5 embryos and in adult cells and the rat umbilical cord mesenchymal stromal cell feeder layer. The methylation of the proximal promoter during development corresponds to what has been observed in mouse em-

bryos [23]. Significant differences were seen between the methylation of ESC-like cells and D4.5 embryos and the other tissues examined (see Fig. (5)). Furthermore, the methylation results correlated with quantitative reverse transcriptase polymerase chain reaction (qRT-PCR). That data showed that Oct4 mRNA is relatively abundant in D4.5 embryos and in ESC-like cells at P3 and P9 and is down-regulated during embryonic development, and in adult tissue. Importantly, these data indicate that rat ESC-like cells express more Oct4 mRNA at passage 3 and 9 than adult liver and brain, or D10.5 and D12.5 embryos, and similar levels of



**Fig. (6).** Quantitative reverse transcription polymerase chain reaction (qRT-PCR) evaluation of Oct4 gene expression. Total RNA samples were collected from D4.5, D10.5, D12.5, Fischer 344 rat adult brain and liver, Fischer 344 rat embryonic stem cell-like cells at passage 3 and passage 9 and from embryonic stem cell-like cells that had been cultured for 7 or 14 days on low-adherence dishes without cytokines or growth factors to form embryoid bodies. Panel A shows results obtained when 4 ng total RNA was used. Panel B shows results obtained when 40 ng total RNA was used. Samples were run in duplicate and averaged on a Bio-Rad iCycler. A control gene was used as the control gene and subtracted from the Ct values of the Oct4 to obtain Delta Ct (see Methods for details). Abbreviations: ESC, rat embryonic stem-like cells; D4.5, embryos collected at 4.5 days post coitus; D10.5, embryos collected at 10.5 days post coitus; D12.5, embryos collected 12.5 days post coitus; EB, F344 rat ESC-like cells were differentiated for 7 or 14 days on low adherence plates in the absence of feeders, cytokines and growth factors; Ct, Threshold cycle: cycle where arbitrary threshold is exceeded (statistic is calculated by Bio-Rad software).



Oct4 mRNA to what is observed in D4.5 embryos. In addition, the differentiation of ESC-like cells to EB resulted in a decreased expression of Oct4 after 7 or 14 days of differentiation.

Regarding the qRT-PCR data, PBGD, a low-abundance housekeeping gene was selected as the internal control. While PBGD has been evaluated as an internal qRT-PCR control in other contexts [48-54], PBGD levels were not evaluated for stability here. This assumption may influence the accuracy of the qRT-PCR results. As reviewed by Lechniak, beta actin, glyceraldehyde-3-phosphate-dehydrogenase and ribosomal RNA have been used for internal controls for qRT-PCR in mammalian oocytes and embryos, and none of these genes is ideal [55]. On the other hand, as reviewed above, there are previous reports that Oct4 gene expression is downregulated during embryonic development and following differentiation of embryonic stem cells to form embryoid bodies. Thus, it is unlikely that the qRT-PCR results are grossly inaccurate.

## CONCLUSION

These results suggest that the architecture of the proximal and distal enhancers of Oct4 and the proximal promoter of Oct4 in the F344 rat is similar to that in mouse, bovine and human. Further, these results indicate that at least in part, the rat Oct4 gene is regulated by DNA methylation during early embryogenesis. With the availability of true rat ESCs, future work comparing the Oct4 regulation machinery between rat and mouse will be enabled [36-38]. Thus, one could investigate, for example, whether orphan nuclear receptor GCNF, a factor shown to act *via* binding the DR0 element located in the Oct4 proximal promoter, plays a role in methylation of the Oct4 gene silencing during embryo development and ESC differentiation in the rat as it does in mouse [56-58].

## ACKNOWLEDGEMENTS

D. Troyer is thanked for reviewing a draft of the Ms. The authors thank B. Lutjemeier, K. Seshareddy, D. Albertini, C. Mauer, N. Bammes, K. Fox, M. Soares, B. Carter, S. Barrett, D. Madison and J. Cox for their assistance. This work was supported by NIH NS34160 and KCALSI to MLW, Howard Hughes undergraduate program in the Division of Biology to MM. The Northwest Kansas Parkinson's Association is thanked for their generous support. Additional support from K-INBRE, Terry C. Johnson Center for Basic Cancer Research, State of Kansas legislature, Dept of Anatomy and Physiology and the Provost' office at Kansas State University. This work is dedicated to David Keefer Weiss and Glenn I Hatton.

## REFERENCES

- [1] Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 2005; 309: 1369-73.
- [2] Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr Biol* 2001; 11: 1553-8.
- [3] Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997; 385: 810-3.
- [4] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; 126: 663-76.
- [5] Takahashi K, Tanabe K, Ohnuki M, *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; 131: 861-72.
- [6] Wernig M, Meissner A, Foreman R, *et al.* *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007; 448: 318-24.
- [7] Yu J, Vodyanik MA, Smuga-Otto K, *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007; 318: 1917-20.
- [8] Pesce M, Scholer HR. Oct-4: gatekeeper in the beginnings of mammalian development. *Stem Cells* 2001; 19: 271-8.
- [9] Nichols J, Zevnik B, Anastassiadis K, *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 1998; 95: 379-91.
- [10] Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 2000; 24: 372-6.
- [11] Caiafa P, Zampieri M. DNA methylation and chromatin structure: the puzzling CpG islands. *J Cell Biochem* 2005; 94: 257-65.
- [12] Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003; 33(Suppl): 245-54.
- [13] Shiota K. DNA methylation profiles of CpG islands for cellular differentiation and development in mammals. *Cytogenet Genome Res* 2004; 105: 325-34.
- [14] Gidekel S, Bergman Y. A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element. *J Biol Chem* 2002; 277: 34521-30.
- [15] Hattori N, Nishino K, Ko YG, *et al.* Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *J Biol Chem* 2004; 279: 17063-9.
- [16] Marikawa Y, Fujita TC, Alarcon VB. Heterogeneous DNA methylation status of the regulatory element of the mouse Oct4 gene in adult somatic cell population. *Cloning Stem Cells* 2005; 7: 8-16.
- [17] Yamazaki Y, Fujita TC, Low EW, Alarcon VB, Yanagimachi R, Marikawa Y. Gradual DNA demethylation of the Oct4 promoter in cloned mouse embryos. *Mol Reprod Dev* 2006; 73: 180-8.
- [18] Okazawa H, Okamoto K, Ishino F, *et al.* The oct3 gene, a gene for an embryonic transcription factor, is controlled by a retinoic acid repressible enhancer. *EMBO J* 1991; 10: 2997-3005.
- [19] Zhang X, Zhang J, Wang T, Esteban MA, Pei D. Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. *J Biol Chem* 2008; 283: 35825-33.
- [20] Yang HM, Do HJ, Kim DK, *et al.* Transcriptional regulation of human Oct4 by steroidogenic factor-1. *J Cell Biochem* 2007; 101: 1198-209.
- [21] Okumura-Nakanishi S, Saito M, Niwa H, Ishikawa F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem* 2005; 280: 5307-17.
- [22] Nordhoff V, Hubner K, Bauer A, Orlova I, Malapetsa A, Scholer HR. Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences. *Mamm Genome* 2001; 12: 309-17.
- [23] Yeom YI, Fuhrmann G, Ovitt CE, *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* 1996; 122: 881-94.
- [24] Yoshimizu T, Sugiyama N, De FM, *et al.* Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev Growth Differ* 1999; 41: 675-84.
- [25] Rosner MH, Vigano MA, Ozato K, *et al.* A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* 1990; 345: 686-92.
- [26] Shi Y, Despots C, Do JT, Hahm HS, Scholer HR, Ding S. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 2008; 3: 568-74.
- [27] Huangfu D, Osafune K, Maehr R, *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* 2008; 26: 1269-75.
- [28] Niwa H. How is pluripotency determined and maintained? *Development* 2007; 134: 635-46.
- [29] Nichols J, Smith A, Buehr M. Rat and mouse epiblasts differ in their capacity to generate extraembryonic endoderm. *Reprod Fertil Dev* 1998; 10: 517-25.
- [30] Buehr M, Nichols J, Stenhouse F, *et al.* Rapid loss of Oct-4 and pluripotency in cultured rodent blastocysts and derivative cell lines. *Biol Reprod* 2003; 68: 222-9.

- [31] Fandrich F, Lin X, Chai GX, *et al.* Preimplantation-stage stem cells induce long-term allogeneic graft acceptance without supplementary host conditioning. *Nat Med* 2002; 8: 171-8.
- [32] Iannaccone P, Taborn G, Garton R. Preimplantation and postimplantation development of rat embryos cloned with cumulus cells and fibroblasts. *Zygote* 2001; 9: 135-43.
- [33] Ouhibi N, Sullivan NF, English J, Colledge WH, Evans MJ, Clarke NJ. Initial culture behaviour of rat blastocysts on selected feeder cell lines. *Mol Reprod Dev* 1995; 40: 311-24.
- [34] Ruhnke M, Ungefroren H, Zehle G, Bader M, Kremer B, Fandrich F. Long-term culture and differentiation of rat embryonic stem cell-like cells into neuronal, glial, endothelial, and hepatic lineages. *Stem Cells* 2003; 21: 428-36.
- [35] Vassilieva S, Guan K, Pich U, Wobus AM. Establishment of SSEA-1- and Oct-4-expressing rat embryonic stem-like cell lines and effects of cytokines of the IL-6 family on clonal growth. *Exp Cell Res* 2000; 258: 361-73.
- [36] Buehr M, Meek S, Blair K, *et al.* Capture of authentic embryonic stem cells from rat blastocysts. *Cell* 2008; 135: 1287-98.
- [37] Li C, Yang Y, Gu J, Ma Y, Jin Y. Derivation and transcriptional profiling analysis of pluripotent stem cell lines from rat blastocysts. *Cell Res* 2009; 19: 173-86.
- [38] Li P, Tong C, Mehrian-Shai R, *et al.* Germline competent embryonic stem cells derived from rat blastocysts. *Cell* 2008; 135: 1299-310.
- [39] Deb-Rinker P, Ly D, Jezierski A, Sikorska M, Walker PR. Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. *J Biol Chem* 2005; 280: 6257-60.
- [40] Boyer LA, Lee TI, Cole MF, *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005; 122: 947-56.
- [41] Minucci S, Botquin V, Yeom YI, *et al.* Retinoic acid-mediated down-regulation of Oct3/4 coincides with the loss of promoter occupancy *in vivo*. *EMBO J* 1996; 15: 888-99.
- [42] Chew JL, Loh YH, Zhang W, *et al.* Reciprocal transcriptional regulation of Pou5f1 and Sox2 *via* the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 2005; 25: 6031-46.
- [43] Pugh BF, Tjian R. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev* 1991; 5: 1935-45.
- [44] Ben-Shushan E, Sharir H, Pikarsky E, Bergman Y. A dynamic balance between ARP-1/COUP-TFII, EAR-3/COUP-TFI, and retinoic acid receptor:retinoid X receptor heterodimers regulates Oct-3/4 expression in embryonal carcinoma cells. *Mol Cell Biol* 1995; 15: 1034-48.
- [45] Pikarsky E, Sharir H, Ben-Shushan E, Bergman Y. Retinoic acid represses Oct-3/4 gene expression through several retinoic acid-responsive elements located in the promoter-enhancer region. *Mol Cell Biol* 1994; 14: 1026-38.
- [46] Sylvester I, Scholer HR. Regulation of the Oct-4 gene by nuclear receptors. *Nucleic Acids Res* 1994; 22: 901-11.
- [47] Yang HM, Do HJ, Oh JH, *et al.* Characterization of putative cis-regulatory elements that control the transcriptional activity of the human Oct4 promoter. *J Cell Biochem* 2005; 96: 821-30.
- [48] Fink L, Stahl U, Ermert L, Kummer W, Seeger W, Bohle RM. Rat porphobilinogen deaminase gene: a pseudogene-free internal standard for laser-assisted cell picking. *Biotechniques* 1999; 26: 510-6.
- [49] Kuhne BS, Oschmann P. Quantitative real-time RT-PCR using hybridization probes and imported standard curves for cytokine gene expression analysis. *Biotechniques* 2002; 33: 1078, 1080-2, 1084.
- [50] Mustajoki S, Kauppinen R, Mustajoki P, Suomalainen A, Peltonen L. Steady-state transcript levels of the porphobilinogen deaminase gene in patients with acute intermittent porphyria. *Genome Res* 1997; 7: 1054-60.
- [51] Raaijmakers MH, van Emst L, de Witte T, Mensink E, Raymakers RA. Quantitative assessment of gene expression in highly purified hematopoietic cells using real-time reverse transcriptase polymerase chain reaction. *Exp Hematol* 2002; 30: 481-7.
- [52] Schmidt U, Fuessel S, Koch R, *et al.* Quantitative multi-gene expression profiling of primary prostate cancer. *Prostate* 2006; 66: 1521-34.
- [53] Thal SC, Wyszkon S, Pieter D, Engelhard K, Werner C. Selection of endogenous control genes for normalization of gene expression analysis after experimental brain trauma in mice. *J Neurotrauma* 2008; 25: 785-94.
- [54] Wang C, Gao D, Vaglenov A, Kaltenboeck B. One-step real-time duplex reverse transcription PCRs simultaneously quantify analyte and housekeeping gene mRNAs. *Biotechniques* 2004; 36: 508-9.
- [55] Lechniak D. Quantitative aspect of gene expression analysis in mammalian oocytes and embryos. *Reprod Biol* 2002; 2: 229-41.
- [56] Fuhrmann G, Chung AC, Jackson KJ, *et al.* Mouse germline restriction of Oct4 expression by germ cell nuclear factor. *Dev Cell* 2001; 1: 377-87.
- [57] Gu P, LeMenuet D, Chung AC, Mancini M, Wheeler DA, Cooney AJ. Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. *Mol Cell Biol* 2005; 25: 8507-19.
- [58] Gu P, Le MD, Chung AC, Cooney AJ. Differential recruitment of methylated CpG binding domains by the orphan receptor GCNF initiates the repression and silencing of Oct4 expression. *Mol Cell Biol* 2006; 26: 9471-83.
- [59] Thiesen HJ, Bach C. Target Detection Assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SP1 protein. *Nucleic Acids Res* 1990; 18: 3203-9.
- [60] van Eijk MJ, van Rooijen MA, Modina S, *et al.* Molecular cloning, genetic mapping, and developmental expression of bovine POU5F1. *Biol Reprod* 1999; 60: 1093-103.
- [61] Takeda J, Seino S, Bell GI. Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic Acids Res* 1992; 20: 4613-20.
- [62] Scholer HR, Dressler GR, Balling R, Rohdewohld H, Gruss P. Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *EMBO J* 1990; 9: 2185-95.